

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 15:09:52 ; Search time 87 Seconds
(without alignments)
435.778 Million cell updates/sec

Title: AAK91826
Perfect score: 965
Sequence: 1 MRGPRSRHGRDAPAPPTPCV.....ATELGSTELVTTKTAGEPQQ 184

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	410.5	42.5	175	11 Q8R4W8	Q8R4W8 mus musculu
2	120	12.4	703	5 Q9NSP3	Q9NSP3 caenorhabd
3	119	12.3	552	10 Q96343	Q96343 brassica na
4	116.5	12.1	3247	12 Q65553	Q65553 bovine herp
5	116	12.0	306	16 Q9EMW2	Q9EMW2 streptomyce
6	116	12.0	862	4 Q9NT23	Q9NT23 homo sapien
7	112	11.6	530	11 Q88276	Q88276 mus musculu
8	111.5	11.6	957	4 Q8TD92	Q8TD92 homo sapien
9	111.5	11.6	991	4 Q9HC15	Q9HC15 homo sapien
10	111	11.5	801	5 Q23635	Q23635 caenorhabd
11	109	11.3	221	16 Q9AD79	Q9AD79 streptomyce
12	109	11.3	3084	12 Q8UZ11	Q8UZ11 pseudorabla
13	108.5	11.2	994	10 Q9ZK75	Q9ZK75 belaginelia
14	108	11.2	1318	12 P90493	P90493 herpes simp
15	106.5	11.0	268	11 Q9CUC7	Q9CUC7 mus musculu
16	106.5	11.0	1212	16 Q9L1C8	Q9L1C8 streptomyce

17	104.5	10.8	2116	12 Q99IE7	Q99IE7 rubella vir
18	104.5	10.8	2116	12 Q99IE5	Q99IE5 rubella vir
19	104.5	10.8	3938	11 Q88778	Q88778 rattus norv
20	104	10.8	283	11 Q9WV47	Q9WV47 mus musculu
21	104	10.8	294	6 Q9MZV0	Q9MZV0 canis famli
22	103.5	10.7	269	5 Q8T6J7	Q8T6J7 clonorchis
23	103.5	10.7	437	6 Q29109	Q29109 sus scrofa
24	103.5	10.7	924	12 Q99307	Q99307 epstein-bar
25	103	10.7	285	5 Q8SXU6	Q8SXU6 dirosophila
26	103	10.7	1611	3 Q42854	Q42854 schizosach
27	103	10.7	1992	3 Q9PBT1	Q9PBT1 neurospora
28	103	10.7	3436	12 Q66666	Q66666 equine herp
29	102.5	10.6	243	4 Q95515	Q95515 homo sapien
30	102.5	10.6	356	11 Q88699	Q88699 mus musculu
31	102.5	10.6	492	13 Q98T94	Q98T94 xenopus lae
32	102.5	10.6	677	4 Q8TCL1	Q8TCL1 homo sapien
33	102.5	10.6	931	4 Q969V6	Q969V6 homo sapien
34	102.5	10.6	934	4 Q9P2B0	Q9P2B0 homo sapien
35	102.5	10.6	967	4 Q96Q00	Q96Q00 homo sapien
36	102.5	10.6	1003	4 Q96DN6	Q96DN6 homo sapien
37	102.5	10.6	1883	4 Q96SC5	Q96SC5 homo sapien
38	102.5	10.6	1905	4 Q96SC5	Q96SC5 homo sapien
39	102	10.6	325	11 Q8VCX3	Q8VCX3 mus musculu
40	102	10.6	405	10 Q9AYM0	Q9AYM0 oryza sativ
41	102	10.6	675	11 Q91YM9	Q91YM9 mus musculu
42	102	10.6	1044	4 Q9P206	Q9P206 homo sapien
43	102	10.6	1123	11 Q9DBD5	Q9DBD5 mus musculu
44	102	10.6	3105	12 Q8UZ19	Q8UZ19 cercopithec
45	101.5	10.5	746	4 Q9BU60	Q9BU60 homo sapien

ALIGNMENTS

RESULT 1
ID Q8R4W8 PRELIMINARY; PRT; 175 AA.
AC Q8R4W8;
DT 01-JUN-2002 (TREMBL) 21, Created)
DT 01-JUN-2002 (TREMBL) 21, Last sequence update)
DT 01-JUN-2002 (TREMBL) 21, Last annotation update)
DE TRAF3 binding protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP
SEQUENCE FROM N.A.
RA Mizuno K., Irie S., Sato T.-A.;
RT "Identification of novel TRAF3 binding protein, T3BP, which increases
cellular F-actin content."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF350257; AAL63914.1; -
SQ SEQUENCE 175 AA; 18846 MW; B64EFP4B52EE93B1 CRC64;

Query Match	Score	410.5	DB 11	Length	175;
Best Local Similarity	56.1%				Pred. No. 9.1e-25;
Matches	101;	Conservative	9;	Mismatches	55;
				Indels	15;
				Gaps	6;
QY	6	RSLSGRDAPAPPTPCVPAECFDLVNRHCVACGLRTPRPKPAAGASPPARIALQPGESVGA	65		
DB	9	RSQRSDSSVPTQCNQTECFDPLVNRHCVSCLEFHT--PDYGHSTLSLEGTALQPGP----	62		
QY	66	GAGEALPLPGLLPAPALIGLALV-LVGLVSRRRQRRLGASAEAPDDDKDA-	123		
DB	63	--GSALRPDVALLVGAPALIGLALVGLVSVLSWWRQ-QLRTAS-----PDSEGVQ	115		
QY	124	PEPLDKVITILSPGISDATAPAPPPGPGTTPGSHVVPATLSTGLVTTKTAGEPQ	183		
DB	116	QESLENVFPSSSETPHASPMPPLKEDADSLPHSHSVVPALELSTGLVTTKTAGEPQ	175		

RESULT 2

Q9NSP3 PRELIMINARY; PRT; 703 AA.
 ID Q9NSP3;
 AC Q9NSP3;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-UN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-UN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical 78.4 kDa protein.
 GN H06104.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN 1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium.
 RL Science 282:2012-2018(1998).
 RL (2)
 (2) SEQUENCE FROM N.A.
 NC STRAIN=BRISTOL N2;
 RA Lactell P., Woessner J., Harrison M.;
 RT "The sequence of *C. elegans* comid H06104."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC006651; AAF39870.2; -
 DR HSSP; 006124; 2SHP.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR003395; PTPC_motif.
 DR InterPro; IPR002965; P-rich_extensn.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR InterPro; IPR000242; TYR_PP.
 DR Pfam; PF00102; Y_phosphatase; 1.
 DR PRINTS; PRO1217; PRICHEXTENSU.
 DR PRINTS; PRO0700; PRTPHPHTASE.
 DR SMART; SM00194; PTPC; 1.
 DR SMART; SM00194; PTPC; 1.
 DR SMART; SM00404; PTPC_motif; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS00055; TYR_PHOSPHATASE_PTP; 1.
 DR Hydroxylase; Hypothetical protein.
 SEQUENCE 703 AA; 78402 MW; DA79207AF8035E9D CRC64;

Query Match 12.4%; Score 120; DB 5; Length 703;
 Best Local Similarity 24.7%; Pred. No. 0.15; 79; Indels 28; Gaps 5;
 Matches 43; Conservative 24; Mismatches 79; Indels 28; Gaps 5;

QY 8 LGRDPAAPTPCVAPACFDLVHVCACGLTRPRKPGAGSSAPARTALQPOESVGA 67
 DB 142 LVARPPPTAPRPPK-----SQSPALPSPSPSPGASGAANAPYTESIKST 193
 QY 68 GEARLPLGLLFGAPLLGLALVLAIVGLVSWRRQRRLRGASSAEPDGDKAPPE 127
 DB 194 SNAATEASMMWV--YIIGGVVILLIVIGIAVFLIMRKSK-PSSSEVPAPSAPEP- 248
 QY 128 DKYIILSPGISDQATAPAMPPEGDPGTPPGHSVPVPATELSTELVTTKTAGP 181
 DB 249 -----KPA-PKPDPKDPKDPKDPVPAKVPSPVIVPIDSTIVP 287

RESULT 3
 Q96343 PRELIMINARY; PRT; 552 AA.
 AC Q96343;
 DT 01-FEB-1997 (Tremblrel. 02, Created)

DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Myosinase-binding protein related protein (fragment).
 OS Brascia napue (page).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosida II; Brassicales; Brassicaceae; Brascia.
 NCBI_TaxID=3708;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=20516 OF SYALOPS KARAT;
 RX MEDLINE=97210758; PubMed=9057822;
 RA Taipaleenau J., Falc A., Ek B., Rask L.;
 RT "Myosinase-binding proteins are derived from a large wound-inducible
 RT and repetitive transcript."
 RL Eur. J Biochem. 243:605-611(1997).
 DR EMBL; U59446; AAC08051.1; -
 DR HSSP; P18674; 1UOT.
 DR InterPro; IPR001229; Jacalin_lectin.
 DR Pfam; PF01419; Jacalin; 3.
 DR NON TER
 FT SEQUENCE 552 AA; 59270 MW; 5FE51A3718FBD0FA CRC64;

Query Match 12.3%; Score 119; DB 10; Length 552;
 Best Local Similarity 30.2%; Pred. No. 0.14;
 Matches 51; Conservative 7; Mismatches 55; Indels 56; Gaps 10;

QY 22 AECFDLVHVCACGLTRPRKPGAGSSAPARTALQPOESVGAAGAPLPLGLLFGA 81
 DB 282 AADIALGNH---FBLRTPAPAPAPGAPAPAP-----PSSHPPAPAP-----A 325
 QY 82 PALGLALVLAIVGLVSWRRQRRLRGASSAEPD-GDKDAPPELKVILLSPGISDA 140
 DB 326 PA-----PGGPRAPAPAGAPRSPAP-----GGPRAP 355
 QY 141 TPAMPPEGDPGTPP---PG---HSPV--VPATELSTELVTTKTAGP 181
 DB 356 PAPA-PAPGCPBPAPGAPGAPGAPAPAPAGTATPAPATTTKIGP 403

RESULT 4
 Q65553 PRELIMINARY; PRT; 3247 AA.
 ID Q65553;
 AC Q65553;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-UN-2002 (Tremblrel. 21, Last annotation update)
 DE UL36.
 GN Bovine herpesvirus 1.
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirinae.
 NCBI_TaxID=10320;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=COOPER;
 RA Schwyzer M., Vlcek C., Lowery D.E., Bello L.J., Meyer G., Misra V.,
 RT Thiry B., Paces V.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE OF 1-179 FROM N.A.
 RC STRAIN=COOPER;
 RA Schwyzer M., Styger D., Vogt B., Lowery D.E., Simard C.,
 RA Labolisiere S., Misra V., Vlcek C., Paces V.;
 RL Vet. Microbiol. 0:0-0(1996).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=COOPER;
 RA Schwyzer M.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN (4)
 RP SEQUENCE OF 1-179 FROM N.A.
 RC STRAIN=JURA;

Query Match	12.1%;	Score 116.5;	DB 12;	Length 3247;
Best Local Similarity	29.6%;	Pred. No. 1.4;		
Matches 55;	Conservative 10;	Mismatches 70;	Indels 51;	Gaps 8;

RESULT 5		
Q9EWE2		
ID	Q9EWE2	PRELIMINARY; PRT; 306 AA.
AC	Q9EWE2;	
	01-MAR-2001	(TREMBLrel. 16, Created)
	01-MAR-2001	(TREMBLrel. 16, Last sequence update)
	01-JUN-2002	(TREMBLrel. 21, Last annotation update)
DE	Hypothetical protein SC04865.	

OS Streptomyces coelicolor
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomyces.
OX NCBI_taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M45.
RA Bentley S.D., Chater K.P., Cerdeno-Tarraga A.-M., Challin G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Baerem A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo Y., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parthill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2) ";
RL Nature 417:141-147(2002).
DR EMBL; AL450450; CAC17653.1; --
KW Hypothetical protein.

OY 8 LRGDAAPFCVPAEFCFDLLVRHCVACGLLRTTRPRPAGS-SSPPAPTAQPOESVAG 66
:
Db 1 MRGAAGGGPADRAGAEA-DMGARHEAGDGRDTGPRPAGVPTAEALLAAALRAESAGTE 59
OY 67 AGEAAL-----PLPGLFGAPALLGIATLVALLVLGVSWR 102

RESULT 6
Q9NT23
ID Q9NT23
PRELIMINARY;
PRT; 862 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypotheoretical 90.3 kDa protein (Fragment).
 GN DKFPA344A1010.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP
 RC
 RC TISSUE=TESTIS;
 RA Duesterhoeft A., Labber J., Mewes H.W., Gaassenhuber J., Wiemann S.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL137579; CAB70821.1; -
 DR HSSP; O07960; IRGP
 DR InterPro; IPR000196; RhOGAP.
 DR Pfam; PF00620; RhOGAP; 1.
 DR SMART; SM00324; RhOGAP; 1.
 KW Hypotheoretical protein.
 FT
 FT NON TER 1
 SEQUENCE 862 AA; 90295 MW; 0DE275EBB8CDB4A76 CRC64;

```

Query Match      12.0%; Score 116; DB 4; Length 862;
Best Local Similarity 27.5%; Pred. No. 0.38;
Matches 60; Conservative 17; Mismatches 77; Indels 64; Gaps 11;

QY          GRDAP--APPCVPAECGFDLIVRHCVACGLLRTRP--KPAGASPADRTALQPES----- 62
              ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db          GDPAPPASPAPPASAPPPRV-----TQQAISPRCPISPASPALDISEPLAVS 377
              ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY          63 -----VGAGAGEA-ALPLPGLLFG---AFALIGLALVALVLGVLSWRRRQRRLRG 110
              :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db          378 VPPAVLELGGAGAPASATPTTPALSPERSLRPHIIPILLRGAELPLVDACQEMCKSLRG 437
              :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY          111 ASSAAPDPGDKADEPLDKVIIISPGISDATAPAMPPEGEDP----- 152
              :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db          438 AQGPLGFDMESP LPP--PLSLRPG---GAP--PPPKNPALLMALALAERAQAQVAEQ 489
              :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY          153 -----GTTPGHNVVP--ATLGSTELVTYTTIAGP 181
              :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db          490 QSQOECGCTPPASQSFPFRSLSLIEVGEPIGTSGCGPP 527
              :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

ID 088276 PRELIMINARY; PRT; 530 AA.
AC 088276;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DE N-acetylglucosamine-6-O-sulfotransferase long form.
GN CHST2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scturognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=WHOLE EMBRYOS;
RX MEDLINE=9830482; PubMed=9712885;
RA Uchimura K., Muramatsu H., Kedomatsu K., Fan Q., Kurosawa N.,
RA Mitsuoka C., Kannagi R., Habuchi O., Muramatsu T.;
RT "Molecular cloning and characterization of an N-acetylglucosamine-6-O-
sulfotransferase.";
RT J. Biol. Chem. 273:22577-22583 (1998).
EMBL: AB011452; BAA32138.1; -;
EMBL: AB011451; BAA32137.1; -;
DR MGD; MGI:189160; Chet2.
DR InterPro; IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransfer; 1.
KW Transferrase.
SQ SEQUENCE 530 AA; 57814 MW; A113E1B735C363EC CRC64;

Query Match 11.6%; Score 112; DB 11; Length 530;
Best Local Similarity 25.9%; Pred. No. 0.47;
Matches 49; Conservative 17; Mismatches 67; Indels 56; Gaps 9;

QY 2 RRGPRSLRGDAPAPPCVPAECFDLVRHCVACGLLRTPRPKAGASSPAPRTALQPOE 61
DB 3 RSSPRALPGALPRPLPAPPA-----AVGRALL-PPWPRRAGRWPAAS----- 44
QY 62 SVGAGAGEALPPLGGLFGAPALL---GLALVALVLVGVSWRRORRLRGASAPAD 118
DB 45 -----PLGMKVFRRKALVLCGYALLVLTMLNLDYKMKHEPQQCN----- 89
QY 119 GDKDAPBPLDKVIIISPGISDATAPAMPPEDEPGTTPPGHS-----VPV--DATELGST 171
DB 90 G-----PLGAAV-----GAAGAGWGRGSPAPAPRAHSHMDRTPYRPPAGVAV 136
QY 172 ELVTTKTAG 180
DB 137 PAAAGSAG 145

RESULT 8
Q8TD92 PRELIMINARY; PRT; 957 AA.
AC 08TD92;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hepatocellular carcinoma-associated protein HCA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Dong X., Chen W.;
RT "Identification of genes in the chromosome X that are differentially
expressed in hepatocellular carcinoma.";
RT Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490507; AAM08354.1; -;
SQ SEQUENCE 957 AA; 103253 MW; E09F9161384CC2B5 CRC64;

Query Match 11.6%; Score 111.5; DB 4; Length 957;
Best Local Similarity 25.6%; Pred. No. 0.94;

Matches 54; Conservative 22; Mismatches 86; Indels 49; Gaps 11;

QY 5 PSLRGDAPAPTPCPVPAECFDLVRHCVACGLLRTPRPKAGASSPAPRTALQPOE 59
DB 37 PADVPESDVPQ-----PSDS-OILGLCASSB-----PSTSVLPISABGPTVPTTISA 87
QY 60 -----QESVAGAGEALPPLP---GLLFGAPALLGALVALVLVGVSWRRORRLRGAS 112
DB 88 SASAGQPTISEPGISVLPTEPBGSTSGPPTISGLCTSVTLA-----SEGRNTRSPPT 143
QY 113 SAAPDGDGDAPEPLDKV--IILSPGISDATAPAMP-----PEEDPGT--- 154
DB 144 SSEEP--STSVPTTASEVPTSLPPTPGSGTSTSVPTAYEGPSTSVPTPDGPTSVL 201
QY 155 -TP---PGHSVPVPATELGSTELVTTKTAGP 181
DB 202 PTPGSGPTSVPLAATGISTSVQATPDGEP 232

RESULT 9
Q9HC15 PRELIMINARY; PRT; 991 AA.
AC 09HC15;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE KIAA1587 protein (Fragment).
GN KIAA1587.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVII. The complete sequences of 100 new cDNA clones from brain which
code for large proteins in vitro.";
RT DNA Res. 7:273-281 (2000).
DR EMBL; AB046807; BAB13413.1; -;
DR InterPro; IPR002190; MAGE.
DR Pfam; PF01454; MAGE; 2.
FT NON TER 1
SQ SEQUENCE 991 AA; 106833 MW; 1DD36363402BAC91 CRC64;

Query Match 11.6%; Score 111.5; DB 4; Length 991;
Best Local Similarity 25.6%; Pred. No. 0.98;
Matches 54; Conservative 22; Mismatches 86; Indels 49; Gaps 11;

QY 5 PSLRGDAPAPTPCPVPAECFDLVRHCVACGLLRTPRPKAGASSPAPRTALQPOE 59
DB 71 PADVPESDVPQ-----PSDS-OILGLCASSB-----PSTSVLPISABGPTVPTTISA 121
QY 60 -----QESVAGAGEALPPLP---GLLFGAPALLGALVALVLVGVSWRRORRLRGAS 112
DB 122 SASAGQPTISEPGISVLPTEPBGSTSGPPTISGLCTSVTLA-----SEGRNTRSPPT 177
QY 113 SAAPDGDGDAPEPLDKV--IILSPGISDATAPAMP-----PEEDPGT--- 154
DB 178 SSEEP--STSVPTTASEVPTSLPPTPGSGTSTSVPTAYEGPSTSVPTPDGPTSVL 235
QY 155 -TP---PGHSVPVPATELGSTELVTTKTAGP 181
DB 236 PTPGSGPTSVPLAATGISTSVQATPDGEP 266

RESULT 10
Q23635 PRELIMINARY; PRT; 801 AA.
AC 023635;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
 DE ZK84.1 protein.
 GN ZK84.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almocough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Cratton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kerhaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Ricken L., Roopra A., Saunders D., Showkhen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterson R.,
 RA Watson A., Weinstock L., Wilkison-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Kirsten J.;
 RT "The sequence of C. elegans cosmid ZK84.";
 RL Submitted (Apr-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U23181; AAC48204.1; -
 SQ SEQUENCE 801 AA; 71123 MW; 070D8F085A71EF28 CRC64;
 Query Match 11.5%; Score 111; DB 5; Length 801;
 Best Local Similarity 25.4%; Pred. No. 0.86;
 Matches 45; Conservative 12; Mismatches 64; Indels 56; Gaps 8;
 QY 12 DAPAPTPCVAECFLLVHCVACGLRTPRPKAGASGPAPRTA-LOPQESVAGAGGA 70
 DB 405 DTPAPASAAAE-----TPAPAPAEETPAPETASAPADAGGAAPADV 448
 QY 71 ALPLPGLFGAPALGLALVALVLVGLVSWRRORRLRGASSAPAPDGDKAP-EPDLDK 129
 DB 449 AAP-ADVATTAE-----TSSAGCAAGSYDVSPAPSE 480
 QY 130 VLI-----LSRGISDAPAPAPPPGGDPTPGHGVPPVATLSTELVTTKTACP 181
 DB 481 VTAPIVESATAPSDSAPADIGPAPSE-PAPAP---IEAPATDAATLETPAPAPAE 532
 RESULT 11
 ID Q9AD79 PRELIMINARY; PRT; 221 AA.
 AC Q9AD79;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE Putative secreted protein (Putative membrane protein).
 GN SCK13.21C OR SCO4929.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxId=1902;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Rendenbach M., Kiese H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kiese H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kiese T., Latke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 DR EMBL; AL512667; CAC21630.2; -
 DR EMBL; AL511182; CAD30919.1; -
 DR InterPro: IPR002965; P. rich exten.
 DR PRINTS; PR01217; PRICHEXTENSN.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 221 AA; 21600 MW; 5EDE3A6C31AFCB89 CRC64;
 Query Match 11.3%; Score 109; DB 16; Length 221;
 Best Local Similarity 33.1%; Pred. No. 0.32;
 Matches 47; Conservative 9; Mismatches 62; Indels 24; Gaps 8;
 QY 39 RTRPKPAGASSAPARTLLOPQESVAG-----AGAPALPLPGLFG---APALLGLA 88
 DB 71 RDPSPDP--ATPAPAGTVGEPS---GAGPSATSWAGLGSPPGLIRGLAPAPATGPA 125
 QY 89 LVLAIV-LVGLVSWRRORRLRGASSAPAPDGD---KAPEDLDKVIILSPISGATAP- 143
 DB 126 VPLPTAPPVRRTPGABETPKPGGAGERERERDDTGERAPAPGTP---QPSRSDRPVPH 181
 QY 144 AMPPPGDDPTPGHGVPPVA 165
 DB 182 PVPPPPPGGTTPPSDPAPEPA 203

RESULT 12
 ID Q8UZ11 PRELIMINARY; PRT; 3084 AA.
 AC Q8UZ11;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE U136 protein.
 OS Pseudorabies virus (strain Kaplan) (PRV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirinae.
 OX NCBI_TaxId=33703;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=KAPLAN;
 RA Klupp B.G., Fuchs W., Granow H., Nixdorf R., Mettenleiter T.C.;
 RT "The Pseudorabies virus U136 tegument protein physically interacts
 RT with the U137 protein."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ422133; CAD19511.1; -
 DR InterPro: IPR005210; Herpes U136.
 DR Pfam; PF03586; Herpes U136; 1.
 SQ SEQUENCE 3084 AA; 324401 MW; ECAD9E1E3DC22D1A CRC64;
 Query Match 11.3%; Score 109; DB 12; Length 3084;

Best Local Similarity 25.7%; Pred. No. 4.9;
Matches 55; Conservative 11; Mismatches 70; Indels 78; Gaps 9;

QY 5 PRELRGDAPAPPCVPAECFDLVRHCVAAGLLRPRKPAACSS-----PAPRTALQPG 60
DB 2791 PPSRPAAPPAPPPSRAPRPE---RHGDEPRRLMTPTPLPPQPSQORPEALWTPEPR 2847
QY 61 ESVC-----AGAGEAALPLPGLLFGAPALLGLVLAIVLVGWSRRR 104
DB 2848 DSAGPPALFTPPSPDPVPGFGEPEPLPP-----PPQPL-----W--- 2885
QY 105 QRLRGASSAEPDGDKDAPEPLDKYIIISPGISDATAAPWPPGDEPGTT---PGHSV 161
DB 2886 -----ASEGEAETPATPLRSPLERM-----PMPPLEAPGTTPPLPPEFAP 2926
QY 162 PVPATELGS-----TELVTTKTAGEPOQ 184
DB 2927 PAPSRLADPHQPPAPWFDLTALVPPAPPPRRQ 2960

ULT 13
QY 92R75 PRELIMINARY; PRT; 994 AA.
AC 092R75;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE SL-TSP/P
OS Selaginella lepidophylla.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Lycopodiophyta; Isoetesopsida; Selaginellales; Selaginellaceae;
OC Selaginella.
RN NCBI_TaxId=59777;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=2.5 HOURS DESIGNATED MICROPHYLUS;
RX MEDLINE=99214479; PubMed=10198107;
RA Zentella R., Mascorro-Gallardo R.O., Van Dijk P., Folch-Mallol J.,
RA Bonini B., Van Vaecq C., Gaxiola R., Covarrubias A.A.,
RA Nieto-Sotelo J., Thevelein J.M., Icuriaga G.,
RT "A Selaginella lepidophylla trehalose-6-phosphate synthase complements
RT growth and stress-tolerance defects in a yeast tps1 mutant.";
RL Plant Physiol. 119:1473-1482(1999).
DR EMBL; U96736; AAD00829.1; -
DR InterPro; IPR001830; GT 20.
DR InterPro; IPR001337; Trehalose_Pase.
DR Pfam; PF00982; Glyco_transf_20; 1.
DR Pfam; PF02358; Trehalose_Pase; 1.
DR SEQUENCE 994 AA; 109359 MW; 83CBCEPA9DDDEIAD CRC64;

Query Match 11.2%; Score 108.5; DB 10; Length 994;
Best Local Similarity 24.8%; Pred. No. 1.7;
Matches 37; Conservative 24; Mismatches 61; Indels 27; Gaps 4;

QY 41 PRPKPAGASSPAPRTALQPOESVAGAGEAALPLPGLLFGAPALLGLVLAIVLVGLVS 100
DB 2 PGYPPSSSTSNKKEAGCGAAGGCGCA-----FELPSLASSRERLV----- 47
QY 101 WRKQRRLRGASSAEPDGDKDAPEPLDKYIIISPGISDATAAPWPPGDEPGTTGG-- 158
DB 48 ---RERQLNRQDEDEDEQQALE-AEBAAVATEVPDAVAATPSLSDPEKISSGRG 103
QY 159 -----HSPVPATELGS TELVTTKTAG 180
DB 104 QRLVVANRLPLSATRKGETEMLENESAG 132

RESULT 14
P90493 PRELIMINARY; PRT; 1318 AA.
AC P90493; 012645; 012646;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE HERPES SIMPLEX virus type 2 (Strain HG52), complete genome.
GN RSL.
OS Herpes simplex virus (type 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxId=10310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HG52;
RX MEDLINE=87111457; PubMed=3027242;
RA McGeoch D.J., Moss H.W., McNab D., Frame M.C.,
RT "DNA sequence and genetic content of the HindIII 1 region in the short
RT unique component of the herpes simplex virus type 2 genome:
RT identification of the gene encoding glycoprotein G, and evolutionary
RT comparisons.";
RL J. Gen. Virol. 68:19-38(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HG52;
RX MEDLINE=90278430; PubMed=2161906;
RA Everett R., Fenwick M.,
RT "Comparative DNA sequence analysis of the host shut-off genes of
RT different strains of herpes simplex virus: type 2 strain HG52 encodes
RT a truncated UL41 product.";
RL J. Gen. Virol. 71:1387-1390(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=HG52;
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.,
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2.";
RL J. Gen. Virol. 72:3057-3075(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=HG52;
RX MEDLINE=92356101; PubMed=1322965;
RA Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.,
RT "A novel herpes simplex virus gene (UL49A) encodes a putative membrane
RT protein with counterparts in other herpesviruses.";
RL J. Gen. Virol. 73:2167-2171(1992).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=HG52;
RA Dolan A.,
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z86099; CAB06701.1; -
DR EMBL; Z86099; CAB06707.1; -
DR InterPro; IPR005205; Herpes_ICP4_C.
DR InterPro; IPR005206; Herpes_ICP4_N.
DR Pfam; PF03585; Herpes_ICP4_C; 1.
DR Pfam; PF03584; Herpes_ICP4_N; 1.
DR SEQUENCE 1318 AA; 135101 MW; 9C0BEAF76897D7B2 CRC64;

Query Match 11.2%; Score 108; DB 12; Length 1318;
Best Local Similarity 27.6%; Pred. No. 2.5;
Matches 54; Conservative 11; Mismatches 67; Indels 64; Gaps 8;

QY 41 PRPKPAGASSP-----APRTALQPOE-----SV 63
DB 674 PRPKPAGAAAPRPAADPRRLAWRLRFVRDALVIMRLRGDLRVAGSEAAVAARAVSL 733
QY 64 GAGAGEALP-LPGLLFGAP-----LLGLVLAIVLVGLVSWRRQRL 108
DB 734 VAGALGPALPRSPRLSSAAAADLLPONOSLRPLADTVAAASLAPASAPPEARER 793
QY 109 RGASSAEPDGDKDAPEPLDKYIIISPGISDATAAPWPPGDEPGTTGSHVPVATEL 168
DB 794 KSPAPRABPG--GAPRPKC-----SRADAPRPAAPPA--GAAPAPPTPPPPRR 842
QY 169 GSTELVTTKTAGEPOQ 184

DB 843 PAA-LTRPAGPDPQ 857

RESULT 15

ID 09CUC7 PRELIMINARY; PRT; 268 AA.

AC 09CUC7;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE Adult male testis cDNA, RIKEN full-length enriched library,

DE clone:493417N04, full insert sequence (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=TESTIS;

MEDLINE=21085660; PubMed=11217851;

Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,

Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

Pleischmann W., Gasterland T., Gissi C., King B., Kochiwa H.,

Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

Schirni L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,

Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,

Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,

Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,

Hayashizaki Y.;

RA "Functional annotation of a full-length mouse cDNA collection.";

RT Nature 409:685-690(2001).

DR EMBL; AK016848; BAB30462.1; -;

DR InterPro; IPR002965; P rich_extensn.

DR PRINTS; PRO1217; PRICEXTENSN.

FT NON TER 1

SQ SEQUENCE 268 AA; 27897 MW; FE8766D010BB9204 CRC64;

Query Match 11.0%; Score 106.5; DB 11; Length 268;

Best Local Similarity 29.8%; Pred. No. 0.62;

Matches 57; Conservative 12; Mismatches 87; Indels 35; Gaps 8;

QY 2 RRGPRSLR--GRDAPRT-PCVPAECFDLVHVCACGLRTPRPKAGASSPAPRTAQ 58

DB 53 REGESYRPAEGCAPRTPRPARARLRPPAAPSPAGSLRPPPPSSAAAA----- 106

QY 59 PDSVAGAGG-AAALPLGLFGAPALGLAIVLALVGLVSWRRRRLRGASSAEP 117

DB 107 ---AAGKGAERTASLPF---TAHRLG-----PTVRSATSERVAPALDVTSTTRG 154

QY 118 DGDKDAPEPLDVIILSPGISDATAPAMPPEDEPCT---TPGHVVPATELGSTE 173

DB 155 GGATAAPGAGRRRLRPAAGVGLAVCPVALPGVSAGLLRHSHKAGH-----GSTEG 205

QY 174 VTTKTAGEEQ 184

DB 206 GLRRPLGPAQ 216

Search completed: June 23, 2003, 15:15:51
Job time : 92 secs

THIS PAGE BLANK (USPTO)